FINAL PUBLISHABLE JRP REPORT

JRP-Contract number	NEW06
JRP short name	TraCIM
JRP full title	Traceability for computationally-intensive metrology
Version numbers of latest contracted Annex Ia and Annex Ib against which the assessment will be made	Annex Ia: V1.0 Annex Ib: V1.0
Period covered (dates)	From 1 June 2012 To 31 May 2015
JRP-Coordinator	
Name, title, organisation	Prof. Alistair Forbes
Tel:	+44 (0)208 943 6348
Email:	alistair.forbes@npl.co.uk
JRP website address	http://www.ptb.de/emrp/1390.html
Other JRP-Partners	
Short name, country	JRP-Partner 1 NPL, United Kingdom JRP-Partner 2 CMI, Czech Republic JRP-Partner 3 INRIM, Italy JRP-Partner 4 PTB, Germany JRP-Partner 5 UM, Slovenia JRP-Partner 6 VSL, Netherlands JRP-Partner 7 Hexagon, Germany JRP-Partner 8 Mitutoyo, Germany JRP-Partner 9 Werth, Germany JRP-Partner 10 Zeiss, Germany
REG-Researcher (associated Home Organisation)	
Researcher name, title	Prof. Paul Scott — Start date: 1 July 2012
(Home organisation Short name, country)	HUD, United Kingdom — Duration: 35 months

Report Status: PU Public

EMRP
European Metrology Research Programme
Programme of EURAMET
The EMRP is jointly funded by the EMRP participating countries within EURAMET and the European Union

REG-Researcher (associated Home Organisation)		
Researcher name, title	Prof. Ulrich Lunze	Start date: 1 July 2012
(Home organisation Short name, country)	WHZ, Germany	Duration: 30 months
REG-Researcher (associated Home Organisation)		
Researcher name, title	Prof. Bernd Muller	Start date: 1 October 2012
(Home organisation Short name, country)	Ostfalia-HAW, Germany	Duration: 32 months
REG-Researcher (associated Home Organisation)		
Researcher name, title	Andrew Galloway	Start date: 1 June 2014
(Home organisation Short name, country)	UoY, United Kingdom	Duration: 12 months

TABLE OF CONTENTS

1 Executive Summary

Introduction

Many areas of metrology rely significantly on software that implements mathematical calculations and it is vital that such software is shown to be operating correctly. The project identified calculations relating to a number of priority metrology applications and developed approaches and associated test data to assess the performance of software implementations of those calculations. An information and communications technology (ICT) system was developed that allows mathematical software to be tested using the internet. A testing service for three of the priority metrology applications is now commercially available.

The Problem

Developers of software that implements mathematical calculations increasingly require or desire that testing of their software be carried out by an independent organisation such as a National Metrology Institute (NMI). There are many benefits of such testing to a software developer. The risks and costs associated with bringing software containing errors to market are reduced. The developer is more easily able to demonstrate that the software meets the requirements of users. The reputation of the developer may be enhanced, providing commercial advantage over rivals operating within the same field.

Problems can arise because the specifications of calculations undertaken by software are often expressed in ambiguous language or are missing altogether. The lack of unambiguous specifications of calculations hinders the process of testing metrology software.

Several national and international initiatives have attempted to address the issue of metrology software testing but progress, though useful, has been sporadic. The software testing process involves a number of steps in which data is passed back and forward between the testing organisation and the developer. While the development of communication via the internet has simplified and speeded up the data exchange process, concerns around commercial sensitivity and the security of data must be addressed.

The project therefore sought to implement a novel approach to testing that includes: the provision of specifications of calculations; the generation of reference data (numerical artefacts) to be used to test software; the design of performance metrics to assess the performance and fitness for purpose of software; and the development of a state-of-the-art ICT system that allows software testing to be implemented securely online, at the point of use.

The Solution

Identification of priority metrology applications

Project partners, including four industrial partners, all manufacturers of coordinate measuring machines, identified the following priority metrology applications for which testing services would be made available: least squares geometric element (LSGE) fitting; and minimum zone geometric element (MZGE) fitting.
The priority metrology applications from the Length metrology area identified by project partners were: evaluation of surface roughness parameters; and least squares fitting of data to a NURBS surface.
From the Chemistry metrology area, the priority metrology applications identified were: least squares fitting of a constant plus exponential decay function to data; least squares fitting of a sum of line profiles to data; and principal component analysis.
From the Electricity and Magnetism metrology area, the application of determining the error vector magnitude of a digital signal was identified.
Finally, a number of "Interdisciplinary" metrology applications were identified: least squares polynomial regression; interlaboratory comparisons; and uncertainty evaluation.

Specifications of computational aims

A procedure was developed that allows a calculation (or computational aim) to be specified in a document in which a number of sections must be completed. The sections include Title, Keywords, Mathematical area, Input parameters, Output parameters, Mathematical model, Properties, References and History (of the specification). This procedure enables the specification of any calculation in any metrology domain in a well-defined and unambiguous way. The web-based, searchable "Computational Aims Database" (www.tracim-cadb.npl.co.uk) was developed to facilitate the storage of documents containing specifications of

calculations. Project partners populated the Computational Aims Database with specifications of calculations for the identified priority applications.
The Computational Aims Database is available for users to populate with additional calculations and will be an important component in future software testing activity.

Reference data
For each calculation, project partners have developed software to generate reference input data and corresponding reference results. Software under test is applied to the reference data and the test results obtained are compared, in an appropriate way, with the reference results to provide an assessment of the performance of the software for that reference pair. Additional information may also be provided to indicate bounds on what claims can be made by, and what can reasonably be expected from, software under test. "Benchmarking" reference pairs have been made freely available to download from the Computational Aims Database, and can be used by a software developer to test their software implementations.

Performance metrics
For each calculation, project partners have developed performance metrics to assess the performance of software under test when applied to a reference pair. Where appropriate, the performance metrics take account of the numerical uncertainty and measurement uncertainty information that accompanies the reference data. Procedures to combine the performance metrics returned for multiple reference data sets to determine a single "figure of merit" figure for the software have also been developed. The performance metrics allow an assessment of the uncertainty contribution associated with the software which can be compared with other uncertainty contributions in an uncertainty budget.

ICT system
Project partners designed and developed the components of the "TraCIM system", an ICT system to allow testing of software to be undertaken using the internet. The TraCIM system is designed to make use of the reference data and performance metrics developed in the previous objectives and has three key constituents: the server; the expert module; and the client module. The server lies at the heart of the TraCIM system and communicates with the other two components. The client module is an interface implemented on the computer of the software developer or user, and is responsible for connecting the software under test with the server. The client module requests reference data from the server and, after processing, returns test results to the server. The expert module provides reference data to the server on demand and undertakes the comparison of reference and test results, using the performance metrics.
Expert modules have been developed for three of the identified priority metrology applications, two from the Length metrology area, LSGE fitting and MZGE fitting and the interdisciplinary application of interlaboratory comparisons. For each application, formats have been defined for XML files that contain the reference data to be provided to the software developer or user, and the test results that are subsequently returned to the server. Guidance on developing client modules for the applications has also been written and serve as a template for using the TraCIM system to test any item of metrology software.

Overall, the TraCIM server developed within the project is available to all NMI partners. Following the end of the project, each NMI partner may use the server to provide its own software testing services, for both the calculations considered within the project (for which partners have generated reference data) or for additional calculations (for which reference data needs to be generated).

Commercial testing services have been developed for three of the metrology applications considered during the project, namely LSGE fitting, MZGE fitting, and interlaboratory comparisons. The first two testing services meet the requirement of industry, in particular, the developers of coordinate measuring machines (CMMs).

Impact

Testing services have been implemented for two applications of importance to the manufacturers of CMMs, namely least squares geometric element fitting and minimum zone geometric element fitting. Shortly after the end of the project, one or both of the testing services had been used by nine organisations. It is anticipated that the testing services will continue to be used by software developers. Impact for those developers includes reduced costs associated with undertaking testing of their products and on dealing with the consequences of bring software containing errors to market. Sales of software that is seen to have been tested by an independent organisation may also increase. More generally, environmental impact will be achieved by improvements in the quality of results returned by software products leading to fewer products being unnecessarily scrapped.

The ICT system was deliberately developed to be as generic as possible, and as such can be easily used to provide testing services for additional calculations, not just from metrology but from any domain. Following the end of the project, it is the intention of some project partners to put in place additional testing services. Additionally, during the project the “TraCIM Association” was formed, with the aim of ensuring that the outputs of the project continue to be utilised by software developers and users, and also providing a foundation for future work and international collaboration in the area of software testing. Membership of the association is open to NMIs and Designated Institutes that have particular interest in software testing.

The project’s outputs have been widely disseminated to the relevant metrology and industrial communities. Project partners have been involved in 13 international conferences. Regular discussions with the project partners from the Length metrology area helped to ensure that the testing service being developed satisfied their expectations, and therefore those of other potential users. Three online testing services have been implemented and are now live, for LSGE fitting, MZGE fitting, and interlaboratory comparisons. The LSGE testing service has already been used by three of the project’s four industrial partners, plus an additional four companies from outside the project team. The MZGE validation service has been used by three of the industrial partners, plus two additional companies.

The TraCIM system was developed with the aim of being easily extended to cover other applications from any area for which the use of reference data is appropriate for software testing. The uptake of project outputs will lead to environmental impacts by improvements in the quality of results returned by software products leading to fewer products being unnecessarily scrapped, as well as to financial impacts by reducing the costs software developers are required to spend on testing their products and on dealing with the consequences of bringing software containing errors to market.

2 Project context, rationale and objectives

Context

Many areas of metrology are increasingly reliant upon mathematical software, i.e., software that implements mathematical calculations. Developers of such mathematical software often require or desire that testing of their software be carried out by an independent organisation such as a National Metrology Institute (NMI). For the software developer, there are numerous benefits of such testing, for example:

- The risks and costs associated with bringing software containing errors to market are reduced.
- The developer is more easily able to demonstrate that the software meets the requirements of users.
- The reputation of the developer may be enhanced, providing commercial advantage over rivals operating within the same field.

While NMIs have independently been undertaking testing of metrology software for many years, a lack of collaboration has meant that no coherent framework for such testing exists. While documentary standards may exist that describe calculations to be undertaken, often little or no guidance is provided regarding how software implementing those calculations is to be tested. Furthermore, specifications of calculations are often:

- Are expressed in ambiguous language, leading to different implementations.
- Involve idealised calculations that could never be performed in practice.
- Are missing altogether.

Such problems only serve to hinder the process of testing metrology software.

A number of national and international initiatives, mainly in Europe and North America, have attempted to address the issue of metrology software testing but progress, though useful, has been sporadic.

Some metrology areas have recognised the need to undertake software testing. In Length metrology, for instance, ISO 10360-6 provides a procedure for evaluating the performance implementing least squares fit to data for simple geometries (lines, planes, spheres, etc.). For other calculations, however, there is often confusion and a proliferation of ad hoc approaches of questionable validity.

A common approach to software testing involves the use of “reference pairs”. Each reference pair is made up of “reference data” (or “reference inputs”) and corresponding “reference results” (or “reference outputs”). The reference pair can be thought of as comprising data for which the corresponding results are “known”. Software under test is applied to reference data to obtain “test results”. The test results are then compared, in an appropriate way, with the reference results to provide an assessment of the performance of the software under test for that reference pair. Repeating this procedure for a “large” number of reference pairs allows an overall assessment to be made of the performance of the software under test.

The process of software testing therefore involves a number of steps in which information (reference data, test results) is passed between the provider of the testing service and the user of that service. Throughout the years, this communication has evolved from physical transportation of the data (for example, stored on tape) between service provider and user, to the use of e-mail to transmit data between the two. The IMERA+ project “Metrology for New Industrial Measurement Technologies”, completed in February 2011, included the development of numerical standards (reference pairs) and procedures for internet-aided software validation (IASV) in the area of 3-dimensional coordinate metrology. The project demonstrated the feasibility of using internet services to run client software on pre-assigned datasets in this application area.

Objectives

The project “Traceability for computationally-intensive metrology” aimed to address the issues discussed above and put in place the components required to be able to deliver testing services for metrology software that implements mathematical calculations.

The project sought to implement a novel approach to testing that includes: the provision of specifications of calculations; the generation of reference data (numerical artefacts) to be used to undertake software testing; the design of performance metrics to assess the performance and fitness for purpose of software; and the development of a state-of-the-art ICT system that allows software testing to be implemented securely online, at the point of use.

The key objectives of the project were:

- To identify at least ten priority metrology applications (i.e., calculation or set of calculations), with particular focus on those from Length metrology and develop a general framework for testing metrology software.
- To develop a procedure to ensure that the specification of a calculation in any metrology domain is correct and unambiguous and, for each priority metrology application, provide specifications of calculations.
- For each calculation, to undertake a mathematical analysis and develop software to generate numerical artefacts, i.e., reference input data and reference results for which quantitative statements about their accuracy have been derived (analogous to calibrating a physical artefact).
- For each calculation, to develop performance metrics that allow test results, i.e., results obtained by applying software under test to the reference input data, to be compared with reference results.
- To develop the framework of an ICT system for the online assessment of software implementing mathematical calculations, and make available testing services for calculations identified as significant in coordinate metrology.

Project partners included four manufacturers of coordinate measuring machines (Hexagon, Mitutoyo, Werth and Zeiss) who could provide invaluable guidance towards the development of the testing services in objective 5.

3 Research results

Objective 1: Identification of priority metrology applications

At the beginning of the project, a number of “priority metrology applications” were identified by the project partners. Each priority metrology application comprises a calculation or set of calculations within a particular metrology area. Together, the priority metrology applications would be the focus of subsequent effort relating to the generation of reference data (Objective 3), the development of performance metrics (Objective 4) and (partially) the development of fully operational testing services (Objective 5).

A very important consideration during the identification process was to include those metrology applications that are of particular interest to the project’s industrial partners – Hexagon, Mitutoyo, Werth and Zeiss, all of which are manufacturers of coordinate measuing machines (CMMs) – so that they could provide input and guidance towards both the generation of appropriate reference data and the requirements of the testing services. Further priority metrology applications were chosen according to several criteria, such as their being commonly encountered within metrology, and the potential challenge in generating reference data. While the majority of the priority metrology applications contain calculations that are specific to one metrology area, three priority metrology applications were identified as being applicable to more than one metrology area (and therefore termed “Interdisciplinary”).

Project partners first compiled a shortlist of potential priority metrology applications. During discussions held at the first partners meeting and subsequently via e-mail, the eleven priority metrology applications listed below were identified as those upon which the project would focus (asterixes signify the two priority metrology applications of interest to the project’s industrial partners).

From the Length metrology area:

- * Least squares geometric element fitting (six calculations: calculation of best-fit 3D line, calculation of best-fit sphere, calculation of best-fit 3D circle, calculation of best-fit plane, calculation of best-fit cone, calculation of best-fit cylinder).
- * Minimum zone geometric element fitting (five calculations: calculation of best-fit 2D line, calculation of best-fit 2D circle, calculation of best-fit plane, calculation of best-fit sphere, calculation of best-fit cylinder).

- Surface roughness (five calculations: application of a Gaussian profile filter, calculation of the amplitude surface texture parameter for a primary profile, calculation of the amplitude surface parameter for a roughness or waviness profile, application of a Gaussian areal filter, calculation of the amplitude surface texture parameter for a surface).
- Least squares fitting of a non-uniform rational basis spline (NURBS) surface to data (one calculation).

From the Chemistry metrology area:

- Least squares fitting of a constant plus exponential decay function to data (one calculation).
- Least squares fitting of a sum of line profiles to data (one calculation).
- Principal component analysis (one calculation).

From the Electricity and Magnetism metrology area:

- Determination of the error vector magnitude of a digital signal (one calculation).

Classified as “Interdisciplinary”:

- Least squares polynomial regression (four calculations: weighted least squares, Gauss-Markov regression, generalised distance regression – no correlation, generalised distance regression - correlation).
- Interlaboratory comparisons (four calculations).
- Uncertainty evaluation (five calculations).

Following discussions, project partners were assigned responsibility for the generation of reference data (Objective 3) for specific priority metrology applications.

Objective 2: Specification of computational aims

An identified problem in the testing of metrology software concerns the poor specification, or even a lack of a specification, of the problem to be solved by the software. An important aspect of the project concerned the development of a framework to allow problems to be specified and stored.

A “computational aim” is required to state *what* problem is to be solved or task is to be executed, and not *how* the problem is to be solved or task is to be executed. Decisions about how the problem is to be solved are the concern of the software developer whose responsibility it is to implement the computational aim.

Figure 1 shows how the specification of a computational aim is the foundation of the assessment of the performance of software.

The specification of the computational aim should be unambiguous, complete, free from contradictions, and independent of the environment, such as hardware and software configurations, in which it is to be implemented. The use of natural language to specify a computational aim can lead to specifications that are verbose and ambiguous. A difficulty of using a programming language to specify a computational aim is that the behaviour of the programme can depend on, for example, the choice of hardware and compiler options.

The project proposed to use the abstract and universal language of mathematics to provide specifications of computational aims. The specification of the computational aim of an underpinning mathematical problem is composed of information contained in the following fields:

- Unique identifier – assigned by the TraCIM system.
- Language – the natural language in which the computational aim is written.
- Title – a title or short statement of the computational aim.
- Keywords – a list of keywords relating to the computational aim.
- Mathematical area – the mathematical area to which the computational aim belongs, based on the “Guide to Available Mathematical Software” (GAMS) index.

- Dependencies – a list of dependencies on other computational aims.
- Input parameters – the parameters that must be assigned in order for the computational aim to be executed. The number of input parameters is specified and then for each input parameter the following mandatory information is provided: Symbol, Description, Type, Shape, Constraints.
- Output parameters – the parameters that are assigned as a result of executing the computational aim. The number of output parameters is specified and then for each output parameter the same mandatory information is provided as for an input parameter.
- Mathematical model – a statement of what problem is intended to be solved or task is intended to be executed.
- Signature – a statement (in the form of a function signature) that is indicative of how software implementing the computational aim would be called.
- Properties – a list of properties of the computational aim. Such properties can be useful as the basis for generating reference pairs.
- References – a list of references to supporting papers, reports, guides and documentary standards.
- Notes – any notes that might help with understanding and implementing the computational aim.
- History – a history of the computational aim. The following information is provided: Date created, Author, Amendment information (Date of amendment, Author of amendment, Summary of amendment).

The computational aim as specified above in intended to be generic and not to relate to a particular metrology area, measuring system or instrument. A "refinement" to a computational aim may also be specified. The refinement gives context to the computational aim in terms of a metrology area, measuring system or instrument, and involves providing additional information. For example, a first refinement may associated dimensions with the quantities involved in specifying the computational aim. A second refinement may associate units and particular sets of values with those quantities.

NPL, with input from the other project partners, developed a means to store specifications of computational aims as PDF documents in which details for the fields listed above are provided.

NPL also developed the "Computational Aims Database", a web-based repository for storing such specifications. The database (available at http://www.tracim-cadb.npl.co.uk/) allows for three categories of user:

- "Search only" – search only users do not require registration to use the database and are only able to query the database for specifications that have been approved.
- "Contributor" – contributors require registration to use the database and can upload specifications of computational aims and request that they be released.
- "Editor" – editors require registration to use the database and can approve and request modifications to specifications of computational aims.

Registration can be requested by contacting the database administrator.

The quality of specifications of computational aims contributed to the Computational Aims Database is controlled by the imposition of a "review and approval" process. A specification uploaded by a Contributor is only released, i.e., made externally visible, after it has been approved by an Editor.

For each identified priority metrology application, specifications of computational aims were written by the project partner "responsible" for that metrology application and uploaded to the Computational Aims Database. They were then reviewed and approved by another project partner before being released.

Following the end of the project, the Computational Aims Database will be available for users to populate with additional calculations and will be an important component in future software testing activity.

Figures 2 and 3 show screenshots of the Computational Aims Database.

Although the "review and approval" process imposed by the Computational Aims Database helps to ensure the quality of specifications of computational aims submitted to the database, it is, of course, possible that an error or ambiguity in a specification may not be spotted by a reviewer, leading to an inaccurate or incorrect specification being released. During the final year of the project, the University of York carried out research on the use of formal methods to bring further rigour to the specification of computational aims. Formal methods allow further analysis of a specification to be carried out, using software tools, helping to identify sources of ambiguity that may not be picked up by a reviewer of the specification.

NPL provided the University of York with training on aspects of metrology, covering traceability, calibration and uncertainty. The University of York subsequently provided project partners with training, during a partners meeting, on the mathematics of software engineering.

A conclusion of the University of York research was that formal methods show promise in the analysis of specifications of computational aims, and it would be very useful to undertake further investigation into the use of formal methods following the end of the project. The involvement of the University of York has allowed a completely new set of skills and capabilities to those possessed by the project partners to be applied to this objective.

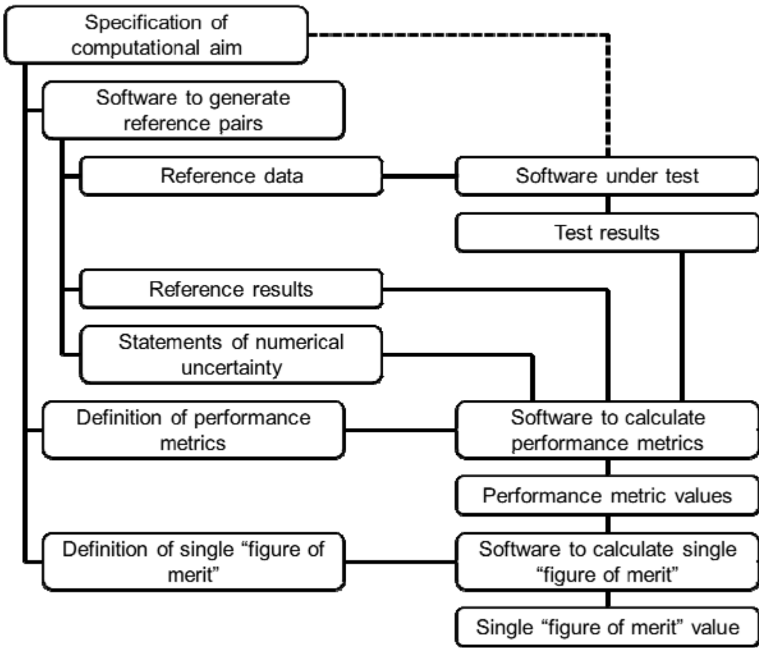

Figure 1: The main steps and components in software testing.

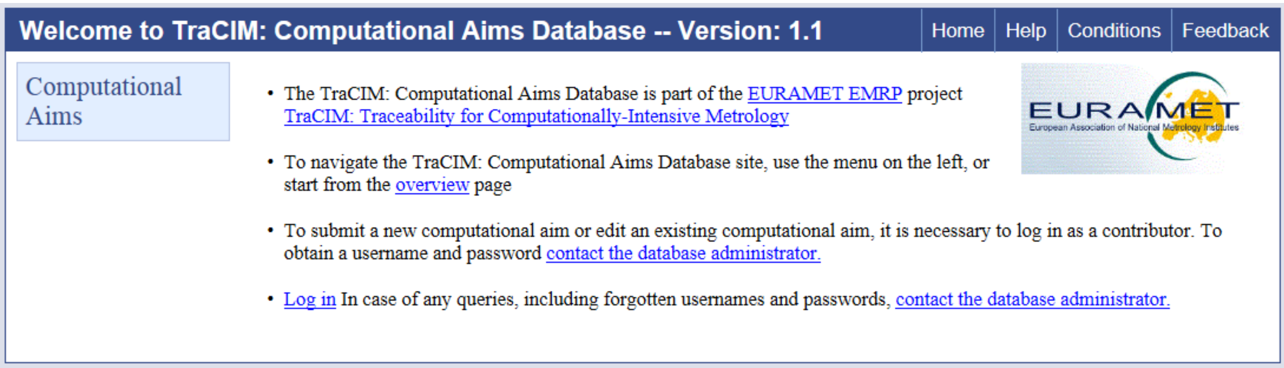

Figure 2: Screenshot of the home page of the Computational Aims Database.

Identifier	Title	Description	Specification	Full Details
en/-/0/000008	Gaussian areal filter	Gaussian filter for calculating surface texture areal parameters.	SurfaceTexture_GaussianArealFilter.pdf [127 KB]	Click for further details
en/-/0/000009	Gaussian profile filter	Gaussian filter for calculating surface texture profile parameters.	SurfaceTexture_GaussianProfileFilter.pdf [110 KB]	Click for further details
en/-/0/000010	Arithmetical mean deviation Pa of assessed profile	Amplitude surface texture parameter for primary profile.	SurfaceTexture_Pa.pdf [105 KB]	Click for further details
en/-/0/000011	Arithmetical mean deviation Ra or Wa of assessed profile	Amplitude surface texture parameter for roughness or waviness profile.	SurfaceTexture_Ra.pdf [112 KB]	Click for further details
en/-/0/000012	Arithmetical mean height of the assessed S–F or S –L surface	Amplitude surface texture parameter Sa for areal S-F or S-L surface.	SurfaceTexture_Sa.pdf [118 KB]	Click for further details
en/-/0/000013	Best fit of a Gaussian line to 3D point coordinates in coordinate metrology	Determine a line to data points in space such, that the sum of the squared distances from a point (xi, yi, zi) to the line is minimized	Gaussian 3D line.pdf [304 KB]	Click for further details
en/-/0/000014	Best fit of a Gaussian sphere to 3D point coordinates in	Determine a sphere to data points such, that the sum of the squared distances from a point	Gaussian Sphere.pdf [303 KB]	Click for further

Figure 3: Screenshot of the Computational Aims Database, listing the specifications of computational aims for a number of the priority metrology applications identified for consideration within the project.

Objective 3: Generation of reference data

Based on the assignment of responsibilities in Objective 1, individual project partners were responsible for developing software to generate reference pairs for particular metrology application areas. However, there were still several opportunities for collaboration between project partners:

- Partner meetings provided a useful forum to discuss potential approaches to data generation, and for each partner to provide an update on progress.
- A researcher from NPL spent three weeks at VSL in 2013 working on data generation approaches.
- The University of Zwickau and PTB collaborated on data generation for the priority metrology application of minimum zone geometric element fitting.
- PTB and UM collaborated on data generation for the priority metrology application of interlaboratory comparisons.

Additionally, at the end of the project, VSL and NPL collaborated on the Metrologia paper “Reference data sets for testing metrology software” [17].

A novel aspect of the data generation to be undertaken within the project was that each reference pair is to be accompanied by information quantifying the numerical properties of the pair, reflecting the fact that, in general in finite precision, reference results are not exactly the solution outputs for the corresponding reference data. Information is required on the numerical accuracy of the reference pair, the numerical sensitivity of the reference results to perturbations in the reference data, and the measurement uncertainty associated with the reference results arising from simulated measurement uncertainty associated with the reference data. This information can be used indicate bounds on what claims can be made by, and what can reasonably be expected from, software under test.

NPL, PTB, VSL, UM, the University of Huddersfield and the University of Zwickau were responsible for generating reference pairs.

For those priority metrology applications involving least squares fitting to data, the null-space method has historically been used to generate reference pairs. This method is an example of an inverse method where one starts with values of the reference results and uses them to generate corresponding values of the reference data. However, the availability of software that implements extended precision arithmetic allows reference software for those problems to be developed and used, in a forward method, to process reference data to obtain corresponding reference results.

Reference pairs have been generated in such a way as to ensure that each reference result is unique to its corresponding reference data, i.e., for a particular reference data set there is only one solution. Therefore, reference pairs have been generated that reflect common practical situations rather than academic examples.

“Benchmarking” reference pairs have been generated made available free to download from the Computational Aims Database (described in the previous section). These reference pairs may be used by a software developer or user to test a software implementation of the corresponding computational aim. “Validation” reference pairs have been generated and are intended to be used as part of testing services for particular calculations. Reference pairs for the priority metrology applications of least squares geometric element fitting, minimum zone geometric element fitting, and interlaboratory comparisons are being used in testing services developed in Objective 5.

The techniques developed during the project, in particular relating to the provision of uncertainty information, may be applied when generating reference data for further calculations.

Objective 4: Development of performance metrics

Performance metrics are used to assign a numerical value to the performance of software when applied to a reference pair. Project partners discussed candidate performance metrics at several partner meetings.

NPL, the University of Huddersfield, and the University of Zwickau developed a number of performance metrics that take into account the “numerical uncertainty” information that accompanies reference pairs. Software to evaluate the performance metrics was developed for the priority metrology applications of minimum zone geometric element fitting, principal component analysis, polynomial regression, and surface roughness.

Candidate approaches for combining performance metrics obtained for a large number of reference pairs to form a single figure of merit for software under test were discussed by PTB, the University of Huddersfield and the University of Zwickau at partner meetings. NPL subsequently developed a short report that summarised the candidate approaches and discussed the advantages and disadvantages of those approaches. NPL and the University of Huddersfield developed software to evaluate a single figure of merit for the priority application areas of principal component analysis, polynomial regression, and surface roughness.

Objective 5: Development of ICT system

PTB, with input from the other project partners, led the design of the ICT system to be used as part of a software testing service. The design considers the communication that must occur between the provider of the testing service and a user of that service. That communication may be summarised as follows:

- The service user must first register to use the testing service.
- The service provider supplies the service user with login credentials (username and password) for the testing service.
- The service user places an order for a test.

- The service provider supplies the service user with an order key.
- The service user requests reference data using the order key.
- The service provider supplies the service user with reference data and a process key.
- The service user submits test results and the process key.
- The service provider providing the service user with a report on the software testing.

It was decided that the first four, administrative, steps would (initially, at least) be undertaken via a combination of e-mail and interaction via a web interface. The final four steps would be undertaken using software, referred to as a TraCIM client (see below).

The University of Ostfalia, with input from PTB and the four industrial partners, developed the state-of-the-art ICT system to be used as the basis for software testing services. The TraCIM client is one of three software modules that the ICT system comprises. The modules are:

- The "TraCIM server" – this software module is a JavaEE application running on a JBoss server. It processes requests (e.g., ordering a test, requesting reference data) from the service user. The TraCIM server is hosted by an NMI.
- The "expert extension" – this software module is incorporated into the TraCIM server and provides reference data to the server on demand and undertakes the comparison of reference and test results. An expert module is required for each calculation to deal with all the logical operations relating to the testing of software implementations of that calculation. The expert extension is also implemented in Java. Importantly, its incorporation into the TraCIM server requires no change to be made to the code of the server.
- The "TraCIM client" – this software module, installed on the PC of the service user, allows the service user to communicate with the TraCIM server using the internet.

Figure 4 provides an illustration of the three modules, and the software under test, within a testing service.

PTB and the University of Ostfalia developed the specification of requirements for, and software modules to implement, secure data exchange within testing services. The software modules were tested by both VSL and the industrial partners.

Expert extensions were developed for the following priority metrology applications:

- Least squares geometric element fitting.
- Minimum zone geometric element fitting.
- Interlaboratory comparisons.

These expert extensions make use of the reference pairs generated as part of Objective 3.

For each of the above three priority metrology application, PTB defined formats for the three XML files that are handled by the expert extension containing:

- The reference data to be supplied to the service user.
- The test results that are submitted by the service user.
- The report on the software testing.

Testing services for the three priority metrology applications have been developed and are fully operational (available at https://tracim.ptb.de/tracim/index.jsf). For each priority metrology application, PTB developed guidance on developing TraCIM clients and that guidance, along with an example TraCIM client, is available on the website of the testing services.

Two workshops, one hosted by PTB, the other by NPL, were held to provide training for project partners and other interested companies in the use of the testing services.

The ICT system has been designed in such a way that testing services for additional calculations can be added straightforwardly. The only requirements for a new testing service, e.g., for a different computational aim, are the development of an expert extension for that computational aim and, potentially, a new TraCIM client. It is intended that further testing services will be made available following the end of the project.

Furthermore, the use of the ICT system is not restricted to testing metrology software – it can be used for mathematical software from any area. It is intended to apply the ICT system to more general mathematical software following the end of the project.

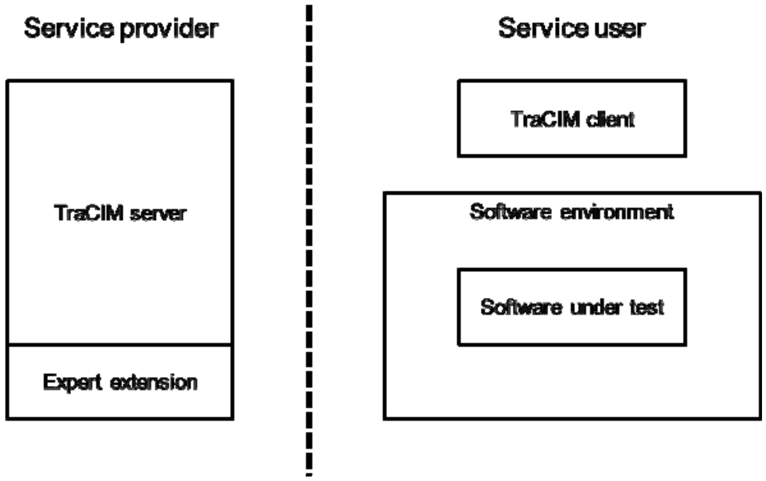

Figure 4: Elements of a testing service, and the software under test.

4 Actual and potential impact

The project's outputs have been widely disseminated to the relevant metrology and industrial communities. Project partners have been involved in 13 international conferences. These conferences include ISPEMI 2014 (the 9th International Symposium on Precision Engineering Measurements and Instrumentation) in Zhangjiajie, China, August 2014, and LMPMI 2014 (the 11th IMEKO Symposium on Laser Metrology for Precision Measurement and Inspection in Industry) in Tsukuba, Japan, September 2014. A session on validation of metrology software was organised at AMCTM 2014 (Advanced Mathematical and Computational Tools in Metrology and Testing), held in St Petersburg, Russia, in September 2014. Project partners also delivered presentations at IMEKO XX World Congress, Busan, Korea, September 2012 and IMEKO XXI World Congress, Prague, Czech Republic, September 2015. In addition to published conference papers, eight articles have been submitted and published in peer-reviewed journals. Project partners have also attended two trade fairs to demonstrate and describe the TraCIM system.

The main element of stakeholder engagement has been achieved through the four unfunded partners involved in the project. These partners represent four major companies involved in the manufacture of coordinate measuring machines, and have a significant interest in ensuring that the software in their machines has been adequately tested. PTB had regular contact with the unfunded partners throughout the project, and their involvement in the project helped to determine the requirements of the ICT infrastructure and the testing services that were developed.

During the project, the "TraCIM Association" was formed with the aim of ensuring that the outputs of the project continue to be utilised by software developers and users, and also providing a foundation for future work and international collaboration in the area of software testing. Membership of the association is open to NMIs and Designated Institutes that have particular interest in software testing.

Regular discussions with the project partners from the Length metrology area helped to ensure that the testing service being developed satisfied their expectations, and therefore those of other potential users.

Towards the end of the project, workshops were held in Germany and the United Kingdom to present, and provide guidance on the use of, the TraCIM system to potential industrial users.

Three online testing services have been implemented and are now live, for LSGE fitting, MZGE fitting, and interlaboratory comparisons. The LSGE testing service has already been used by three of the project's four industrial partners, plus an additional four companies from outside the project team. The MZGE validation service has been used by three of the industrial partners, plus two additional companies. The testing services put in place during the project will continue to be available and it is anticipated that they will be used by more manufacturers of coordinate measuring machines.

Furthermore, the ICT infrastructure developed during the project is sufficiently generic to allow testing services to be developed for additional calculations. Project partners intend to use reference pairs generated within the project to make available testing services for a number of the priority metrology applications. Partners intend to engage with software developers from outside the metrology domain to demonstrate that the outputs of the project are applicable to a wider audience.

During the project, the "TraCIM Association" has been formed, having the following aims:

- Promotion of European and international scientific cooperation in the field of traceability of mathematical evaluation algorithms in metrology.
- Promotion of technical and scientific findings in the field of traceability of evaluation algorithms and software used in industry.
- Transfer of scientific knowledge and information, among other things, for the international harmonisation and recognition of metrological reference algorithms and reference data of the individual National Metrology Institutes and Designated Institutes.
- The promotion of scientific work produced by the members in the field of traceability of mathematical evaluation algorithms in metrology.
- The further development and assurance of quality standards in the case of the validation and traceability of metrological evaluation algorithms.

Membership of the TraCIM Association is open to NMIs, designated institutes (DIs) and individuals with interests in the field of software testing. Activity within the association will help to ensure that the outputs of the project continue to be used, and that the importance of testing of metrology software is further emphasised.

The TraCIM system was developed with the aim of being easily extended to cover other applications from any area for which the use of reference data is appropriate for software testing. It is intended that project partners will make available additional testing services, both for some of the calculations considered within the project and for further calculations. Work is continuing in the field of coordinate metrology, with reference data currently being generated for calculations implemented by involute gear software and hole fitting software. Partners are also engaging with a leading supplier of test and measurement software development tools to make available testing services for a number of the calculations, including least squares polynomial fitting and principal component analysis, considered within the project. Both activities ensure that the outputs of the project, i.e., reference data and the ICT infrastructure, will continue to make impact following the end of the project.

The uptake of project outputs will lead to environmental and financial impacts. Environmental impact will be achieved by improvements in the quality of results returned by software products leading to fewer products being unnecessarily scrapped. Financial impact will be achieved by reducing the costs software developers are required to spend on testing their products and on dealing with the consequences of bringing software containing errors to market. Sales of software that is seen to have been tested by an independent organisation may also increase.

5 Website address and contact details

The TraCIM public website is at http://www.ptb.de/emrp/1390.html.

For general questions about the project please contact Alistair Forbes: alistair.forbes@npl.co.uk

6 List of publications

[1] A B Forbes, H D Minh, Generation of numerical artefacts for geometric form and tolerance assessment, International Journal of Metrology and Quality Engineering, Volume 3, Number 3, 2012.

[2] A B Forbes, Uncertainty associated with form assessment in coordinate metrology, International Journal of Metrology and Quality Engineering, Volume 4, Number 1, 2013.

[3] U Lunze, F Härtig, T Spliedt, D Hutzschenreuter, Generierung von Testdaten zur Auswertung von Geometrieelementen bei Koordinatenmessungen, Wissenschaftliche Schriften des Institutes für Produktionstechnik, Produktionstechnik - innovativ und interdisziplinär, 4. Symposium, Westsächsische Hochschule Zwickau (ISSN 1863-1916).

[4] B Müller, B Wickner, Wiederholbare und nachverfolgbare Zertifizierungstests mit Open-Source – Heute und in 20 Jahren, Proc. Informatik 2013

[5] B Acko, B Sluban, T Tasic, S Brezovnik, Performance metrics for testing statistical calculations in interlaboratory comparisons, Advances in Production Engineering & Management, Volume 9, Number 1, 2014.

[6] B Müller, Repeatable and Tracable Software Verification for 3D Coordinate Measuring Machines, Proc. MEI 2014 (International Symposium on Management, Engineering and Informatics), 2014.

[7] A B Forbes, I M Smith, P M Harris, F Härtig, K Wendt, U Lunze, A new approach to CMM software validation, Laser Metrology and Machine Performance XI, 2015.

[8] A B Forbes, I M Smith, F Härtig, K Wendt, Overview of EMRP Joint Research Project NEW06 "Traceability for computationally-intensive metrology", Advanced Mathematical and Computational Tools in Metrology and Testing X, 2014.

[9] H D Minh, I M Smith, A B Forbes, Determination of numerical uncertainty associated with numerical artefacts for validating coordinate metrology software, Advanced Mathematical and Computational Tools in Metrology and Testing X, 2014.

[10] K. Wendt, M. Franke, F. Härtig, Validation of CMM evaluation software using TraCIM, Advanced Mathematical and Computational Tools in Metrology and Testing X, 2014.

[11] F. Keller, K. Wendt, F. Härtig, Estimation of Test Uncertainty for TraCIM Reference Pairs, Advanced Mathematical and Computational Tools in Metrology and Testing X, 2014.

[12] G J P Kok, I M Smith, Approaches for assigning numerical uncertainty to reference data pairs for software validation, Advanced Mathematical and Computational Tools in Metrology and Testing X, 2014.

[13] B. Acko, S. Brezovnik, L. Crepinsek-Lipus, R. Klobucar, Verification of statistical calculations in interlaboratory comparisons by simulating input datasets, International journal of simulation modelling, Volume 14, Number 2, 2015.

[14] F Härtig, J Tang, D Hutzschenreuter, K Wendt, K Kniel, Z Shi, Online validation of comparison algorithms using the TraCIM system, International Journal of Mechanical Engineering and Automation, Volume 2, Number 7, 2015.

[15] D Hutzschenreuter, F Härtig, K Wendt, U Lunze, H Löwe, Online validation of Chebyshev geometric element algorithms using the TraCIM system, Journal of Mechanical Engineering and Automation, Volume 5, Number 3, 2015.

[16] B Müller, Automatization of acceptance tests for metrology algorithms, Proc. 13th International Conference on Software Engineering Research and Practice, 2015.

[17] G J P Kok, P M Harris, I M Smith, A B Forbes, Reference data sets for testing metrology software, Metrologia, Volume 53, Number 4, 2016.